УДК 004.318

В.П. Бурдаев
Харьковский национальный экономический университет, г. Харьков
Украина, 61001, г. Харьков, пр. Ленина, 9а

Формирование правил базы знаний для функциональной системы

V.P. Burdayev
Kharkov National Economic University, c. Kharkov
Ukraine, 61001, c. Kharkov, Lenin's avenue, 9a,

Rules Formation of the Knowledge Base for Functional System

В.П. Бурдаєв
Харківський національний економічний університет, м. Харків
Україна, 61001, м. Харків, пр. Леніна, 9а

Формування правил бази знань для функціональної системи

В статье рассматриваются различные модели баз знаний для функциональной системы, реализованной в системе «КАРКАС»: модель базы знаний для определения риска возникновения ишемической болезни сердца, модель базы знаний для определения компетентности по информатике и модель возмущения правил базы знаний.
Ключевые слова: модели баз знаний, функциональная система, возмущения правил базы знаний.

In the article, various models of knowledge bases for the functional system realized in system “KARKAS” are considered, i.e. model of the knowledge base for definition of risk of occurrence of an ischemic heart trouble, model of the knowledge base for competence definition on computer science and model of indignation of rules of the knowledge base.
Key Words: models of knowledge bases, functional system, indignations of rules of the knowledge base.

Розглядаються різні моделі баз знань для функціональної системи, реалізованої в системі «КАРКАС»: модель бази знань для визначення ризику виникнення ішемічної хвороби серця, модель бази знань для визначення компетентності з інформатики й модель збурювання правил бази знань.
Ключові слова: моделі баз знань, функціональна система, збурювання правил бази знань.

Введение

Современное развитие Интернета рассматривает распределенные интеллектуальные системы как качественно новые технологии, особенностями которых являются моделирование функциональных систем [1], использование динамически развивающейся онтологии предметной области, мультиагентный подбор адаптивной стратегии принятия решения.

Последние десятилетия особый интерес у исследователей искусственного интеллекта вызывают онтологии. Онтологии могут использоваться для представления знаний о понятиях предметной области (ПрО) и отношениях между ними, а также для описания содержания Web-страниц. Кроме того онтологии можно использовать при построении баз знаний (БЗ) не только экспертных систем (ЭС), но также любых других приложений.

В области интеллектуальных систем одной из основных задач является построение модели онтологии, адекватно отображающей процессы, протекающие в предметной области. На практике возникает ряд проблем в построении адекватной модели, например, предметная область обладает структурой и процессами, которые не могут быть закончены в ограниченный период времени.

Целью данной работы является разработка моделей баз знаний для иерархической функциональной системы, которая адекватна предметной области, обладающей динамической структурой.

Формирование базы знаний в системе «КАРКАС»

Модель онтологии предметной области в системе «КАРКАС» [2-4] состоит из иерархии классов предметной области, связей между ними (правил вывода), которые действуют в рамках этой модели. В системе предложен механизм интерпретации онтологии в условиях динамического изменения ее параметров (базового класса, связей между классами и взаимодействия объектов классов). Техническая реализация онтологии рассмотрена в [4]. Система построена по модульному принципу и по этой причине имеет возможность подсоединения других дополнительных модулей. В архитектуре системы можно выделить следующие основные модули: загрузчик; модуль для разработки БЗ; модуль консультации; модуль кластеризации данных.

Загрузчик осуществляет запуск системы и координирует взаимодействие всех модулей.

Основным компонентом модуля для разработки БЗ является редактор БЗ.

Модуль консультации содержит следующие компоненты: машину вывода, блок объяснения, анализатор тестов, доску объявления, монитор базы фактов.

Модуль кластеризации данных позволяет осуществлять интерактивную и интеллектуальную классификацию многомерных данных.

Модуль Web-сервисов позволяет осуществлять обучение и поиск информации с помощью Web-технологий.

Для тестирования знаний используются генератор тестов и монитор преподавателя.

Модули системы представлены в виде программных пассивных агентов (взаимодействие их осуществляется с помощью массивов параметров и через общую память компьютера). Такие компоненты, как редактор БЗ, машина логического вывода, блок объяснения, являются общими для перечисленных модулей и выступают инструментами для функционирования модулей. В системе выделены два режима использования: режим когнитолога – специалист, моделирующий онтологию ПрО, и режим проблемного специалиста.

Модель функциональной системы для определения риска ИБС

Функциональная система, по П.К. Анохину (ФС), – это система, сформированная для достижения заданного полезного результата (целевой функции) в процессе своего функционирования [1]. Следовательно, ее системообразующим фактором является конкретный результат ФС. Другими словами, цель рассматривается как заданный результат, а ограничения – как степень свободы, необходимая для достижения результата.

Модель предметной области рассматривается как ФС, в которой результат оказывает организующее влияние на все этапы формирования онтологии. Классы и связи между ними можно рассматривать как логическую конструкцию ФС.

Отличительная особенность ФС – в их открытости, неавтономности, неизолированности от внешней среды. Математической моделью, описывающей эволюцию таких систем, служат неавтономные дифференциальные уравнения. Понятие устойчивости таких систем рассмотрены в [4]. В системе модель ФС реализована на основе использования онтологии предметной области.

ФС можно рассматривать, например, как совокупность функций с некоторым набором операций, применяемых к этим функциям. Роль функций играют правила БЗ, а основные операции – это сопоставление атрибута с образцом и определение условий применимости правил.

В системе «КАРКАС» ФС является формализованным отражением предметной области в виде иерархической структуры набора управляющих компонент, которые взаимодействуют между собой для достижения главной цели.

С другой стороны, ФС можно рассматривать как систему высказываний с логическими операциями над ними. Важной особенностью ФС является ее контент, моделирующий предметную область.

В качестве примера рассмотрим построение ФС для определения риска возникновения ишемической болезни сердца (ИБС). Актуальность разработки ее объясняется:

– во-первых, эволюционным переходом от терапии ИБС к профилактике ИБС; от популяционной профилактики к профилактике индивидуальной;

– во-вторых, стремлением человека к самостоятельному получению знаний по оценке риска ИБС;

– в-третьих, возможностью самоконтроля за изменением риска ИБС для принятия решения обращения за врачебной консультацией при высокой ее степени.

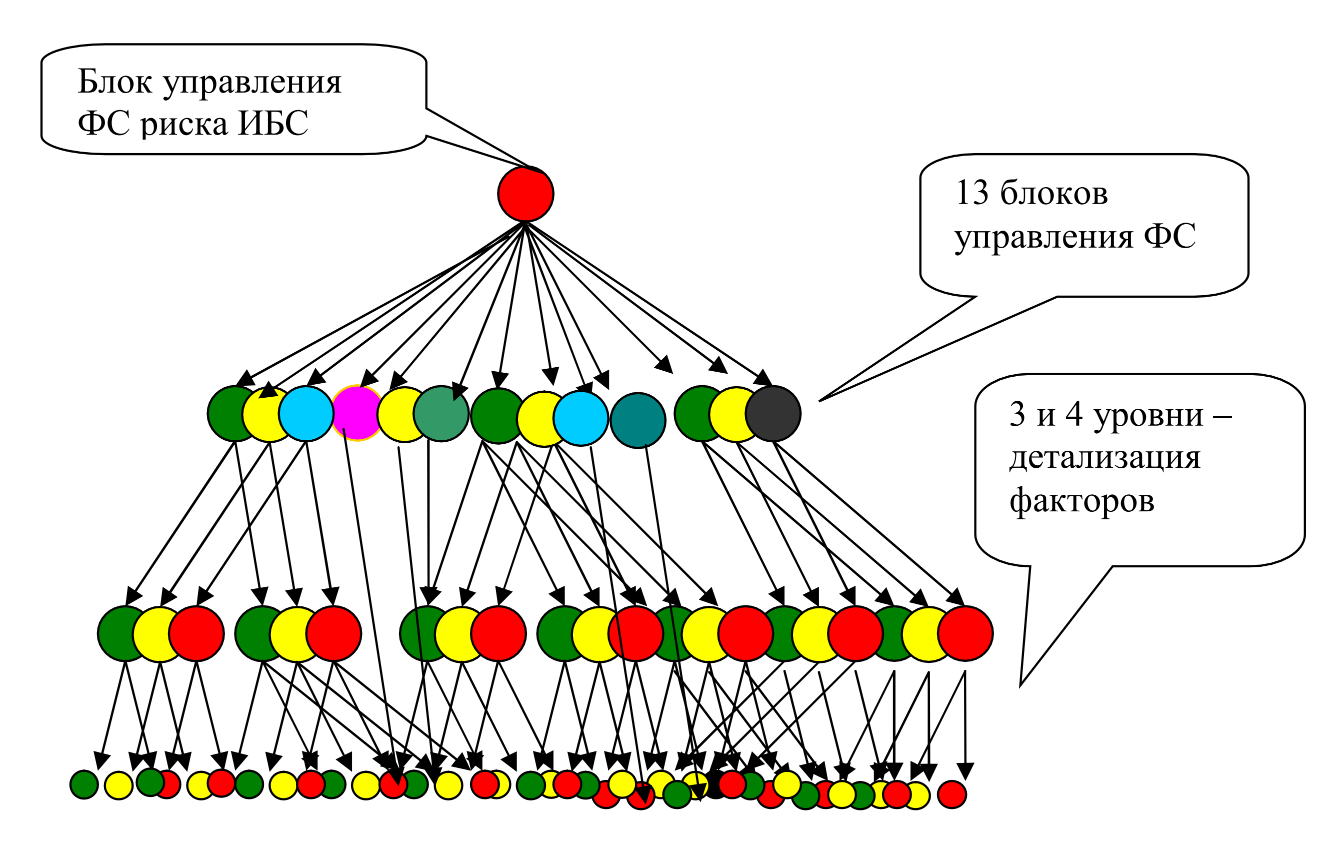


Рисунок 1 – Иерархическая модель ФС для определения риска ИБС

Под фактором риска понимается социальный, биологический и экономический статус человека, модели его поведения и условия, содействующие возникновению ИБС. По результатам исследований экспертов установлено, что основные риск-факторы ИБС:

гиперхолестеринемия, гипертензия, курение, гиподинамия, избыточная масса тела, коронарное поведение, стресс, отсутствие социально-психологической поддержки, нерациональное питание, сахарный диабет, неблагоприятная наследственность, неадекватный отдых, алкоголь.

Таким образом, для оценки риска возникновения ИБС у практически здорового человека отобрано 13 ведущих факторов согласно мнениям экспертов. Эти факторы являются системообразующими для соответствующих 13 ФС, которые взаимодействуют между собой и управляются из основного блока управления функциональной системы риска ИБС. Модель иерархической ФС риска ИБС представлена на рис. 1. Системообразующим фактором ФС является риск ИБС.

На рис. 2 представлено дерево целей иерархической ФС для определения риска ИБС в системе «КАРКАС» в режиме редактирования БЗ.

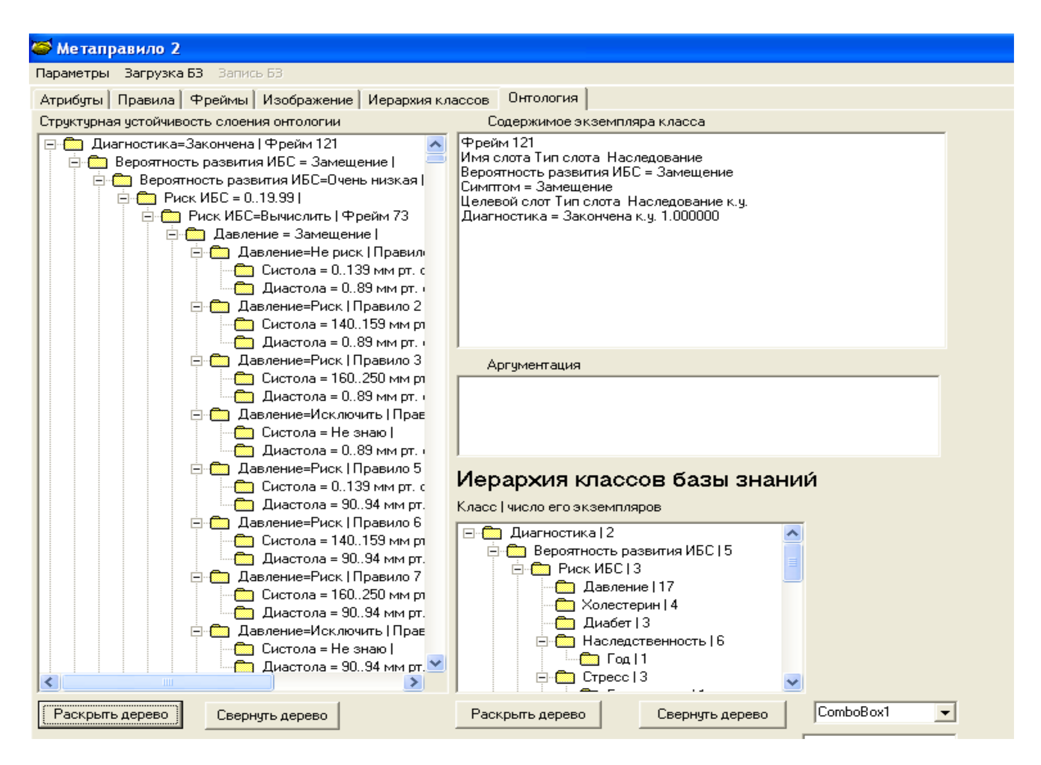


Рисунок 2 – Дерево целей ФС для определения ИБС (режим редактирования БЗ)

Онтология ПрО содержит 13 классов и трехуровневую структуру классов. В БЗ осуществлена кластеризация правил, что позволило улучшить структуру онтологии.

Модель функциональной системы для определения компетентности

Конкурентоспособность на рынке труда определяется профессиональной компетентностью, которая формируется и развивается благодаря изучению современных компьютерных дисциплин. Компетенция – это совокупность поведения, являющаяся результатом знаний, умений, навыков и личностных качеств студента, которые необходимы ему для успешного выполнения необходимых задач. Компетентность – это владение студентом соответствующей компетенцией, содержащей его личностное отношение к предмету деятельности.

Функциональная система компетентности (ФСК) – это система, сформированная для достижения заданного полезного результата (компетентности) в процессе своего функционирования. Системообразующим фактором функционирования ФСК является

конкретный результат – компетентность. Другими словами, цель рассматривается как заданный результат, а ограничения – как степень свободы, необходимая для достижения результата.

Физическая модель БЗ хранит экземпляры классов, объектов, значения атрибутов объектов и логические связи между классами, объектами.

Основные методы оценки компетенций:

– тестирование (психологическое, профессиональное);

– анализ психофизиологических особенностей студентов;

– аттестация (модули, зачеты, экзамены);

– экспертные оценки интеллектуальных и практических компетентностей.

Разработка модели компетентности в системе «КАРКАС» предполагает следующие этапы:

1. Построение онтологии предметной области.

2. Квалиметрия (энтропия) интеллектуальной и практичной компетентностей.

3. Тестирование онтологии.

Модель компетентности – это перечень компетенций с конкретными показателями их проявлений в профессиональной деятельности. В модель включаются компетенции, наиболее важные для студента по информатике.

Эффективная модель предполагает разработку профилей компетенций – наборов компетенций для разных уровней изучения информатики в вузе (например, первый курс финансового факультета – ЗУН (знания, умения, навыки) по MS Office и Интернет-технологиям, второй курс финансового факультета – ЗУН по СУБД Access и MySql и Интернет-технологиям и т.д.). Пример иерархической модели ФСК по информатике, в которой выделены функциональные подсистемы, определяющие соответствующие ключевые компетенции, приведен на рис. 3.

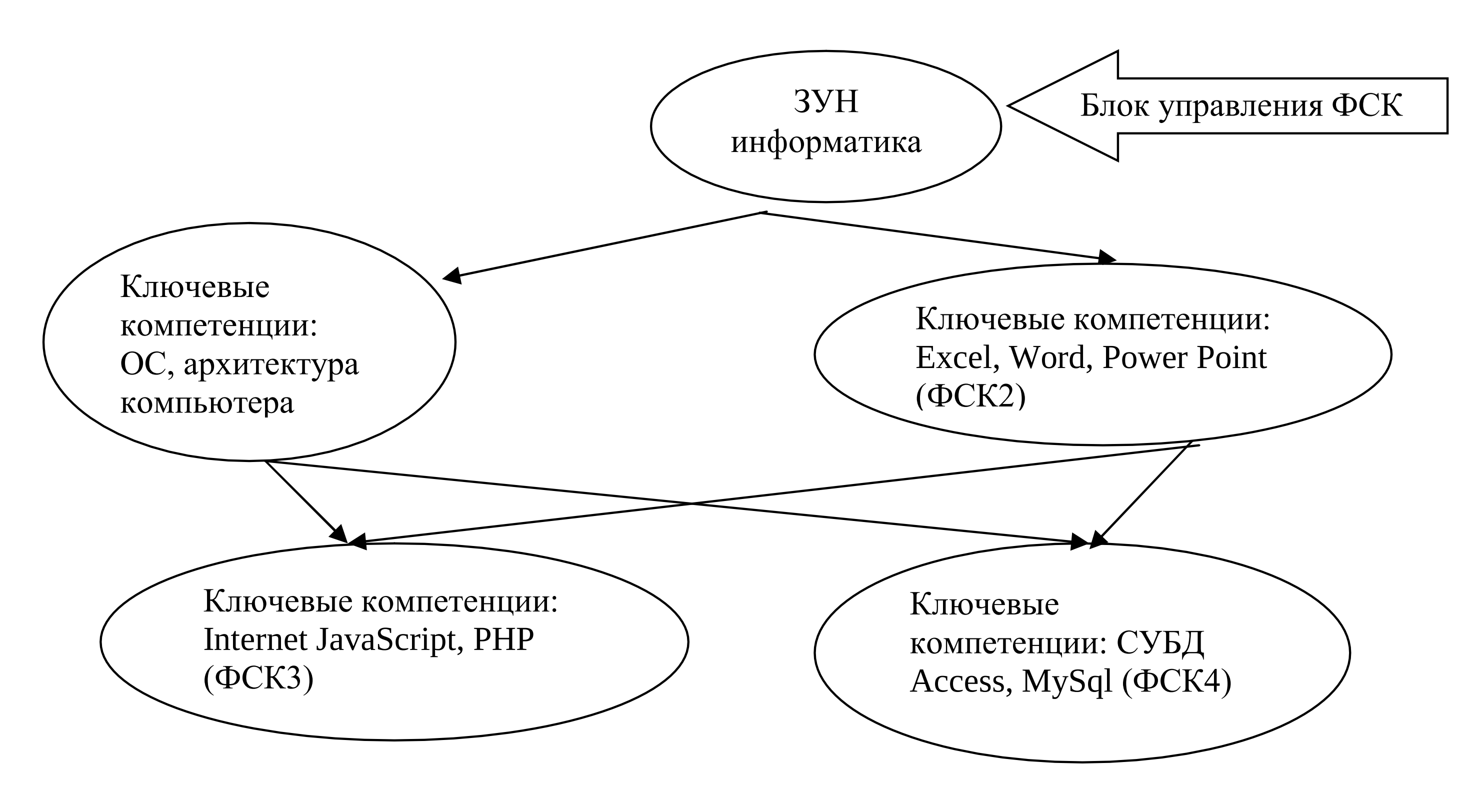


Рисунок 3 – Модель формирования компетентности студента по информатике в системе «КАРКАС»

Алгоритм построения ФСК в системе «КАРКАС».

1. Описать компетенции, необходимые студентам для предметной области.

2. Разработать квалиметрию интеллектуальной и практичной компетенций.

3. Проанализировать эффективность квалиметрии (при необходимости внести корректировки).

4. Выделить ключевые компетенции.

5. Построить модель компетенции по предметной области:

– определить компетенции для каждого раздела дисциплины;

– построить орграф связей компетенций.

6. Разработать методы оценки компетенции (интервью, тесты, ситуационные задачи, бизнес-симуляции).

7. Построить онтологию предметной области (ввести атрибуты, объекты, установить отношения между объектами (написать правила и фреймы).

8. Произвести тестирование базы знаний.

Для ФСК по MS Excel выделены следующие ключевые компетенции:

1. Интерфейс Excel; (x_1).

2. Обработка ввода данных; (x_2).

3. Форматирование; (x_3)

4. Формулы, имена ячеек, ссылки; (x_4).

5. Диаграммы; (x_5).

6. Макросы, сводные таблицы; (x_6).

7. Консолидация; (x_7).

8. Разное. (x_8).

Число возможных логически решений для качественной оценки степени компетентности равно 280 = 8 (объекты) × 5 (уровень компетентности) × 7 (степень компетентности). Тест для оцифровки компетенций содержит свыше 300 вопросов.

На рис. 4 приведено дерево целей ФСК по MS Excel.

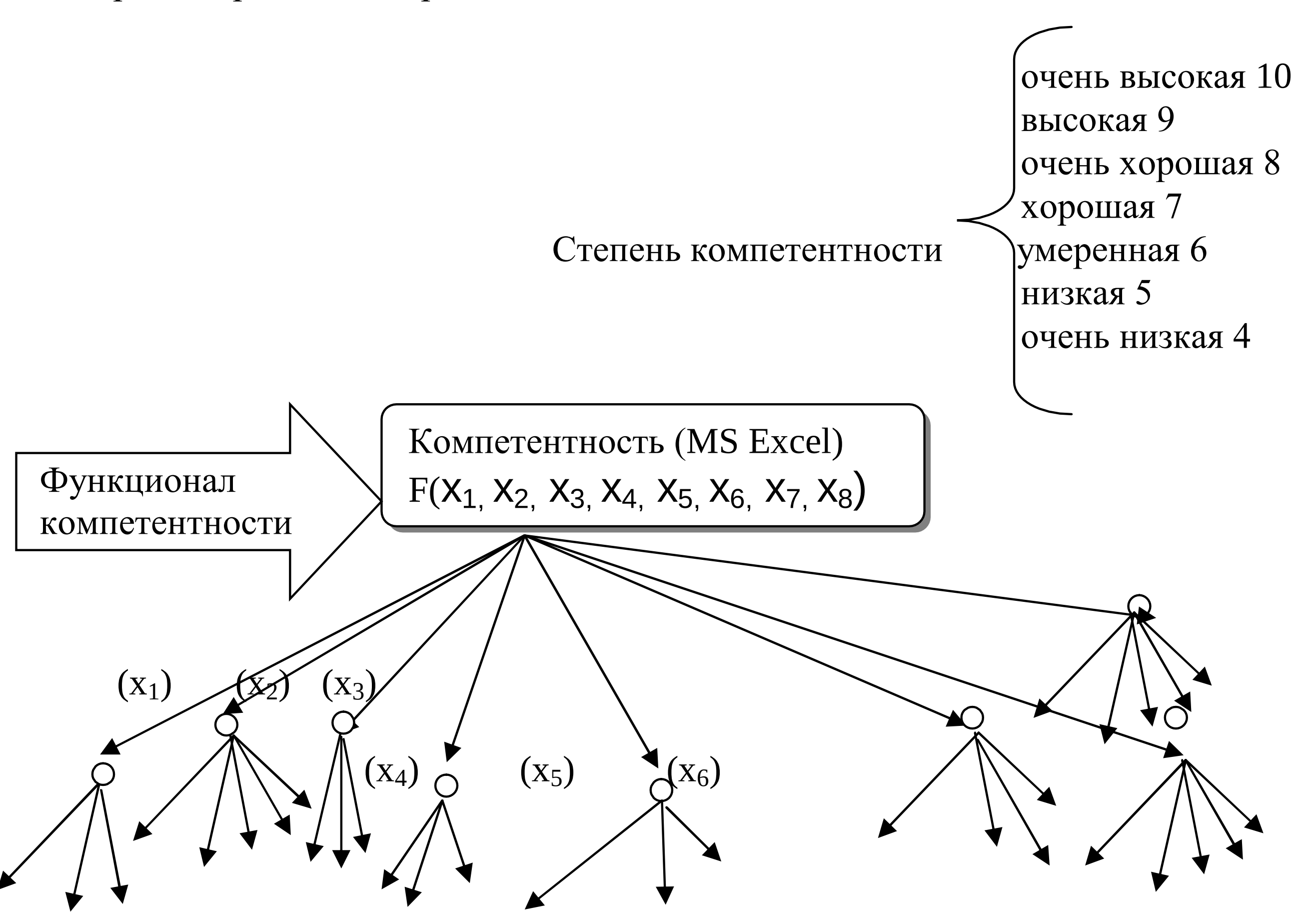


Рисунок 4 – ФСК по MS Excel (дерево целей)

Основные характеристики модулей обучения и тестирования системы «КАРКАС»

Программная реализация системы «КАРКАС» для обучения и тестирования основана на использовании клиент-серверной технологии на основе программирования сокетов. При этом система реализует клиент-серверное взаимодействие: толстый клиент – тонкий сервер, то есть серверная часть реализует только доступ к ресурсам системы (идентификация ученика, установление связи с ним и получение оценок теста от него), а основная часть приложения (машина вывода, объяснение, база знаний) находится на клиенте.

Система «КАРКАС» позволяет эффективно:

– создавать тесты;

– проводить тестирование как на отдельном компьютере, так и по локальной сети;

– составлять по каждому тестированию детальный протокол и осуществлять анализ результатов тестирования;

– формировать ведомости результатов тестирования, которые могут быть использованы для хранения информации об аттестации, блочном контроле, экзамене;

– работать в интеграции с пакетом Microsoft® Office (Word, Excel, PowerPoint);

– использовать подсказку и обучающие блоки по работе с ней (презентации в стиле MS PowerPoint);

– использовать технологию Microsoft Agent и Microsoft Speech API для сопровождения тестирования и во время обучения контента;

– настраивать индивидуальные стратегии для тестирования: выбор тем; формирование тестов по темам; использование датчика случайных чисел для тестов; использование коэффициентов значимости вопросов;

– адаптировать тест по уровню знаний во время тестирования (как в сторону повышения значимости вопросов, так и в противоположную);

– интерпретировать результаты тестирования с помощью диаграммы текущей оценки; диаграммы распределения верных и неверных ответов; диаграммы статистики ответов, модифицированных лиц Чернова;

– формировать тесты динамически согласно стратегиям преподавателя и правилам БЗ.

Для создания теста вопрос имеет следующую структуру.

Название вопроса (атрибут):

– для украинского языка ключевое слово Питання;

– для русского языка ключевое слово Вопрос.

Название вопроса начинается с ключевого слова, а далее может идти произвольный текст. Текст вопроса набирается по желанию пользователя или копируется из буфера обмена. Ответы набираются в окне, расположенном ниже окна ответа. В одной строке ответа может располагаться до 255 символов. Количество ответов на один вопрос может быть до 99.

Текст ответа набирается в первом столбце. Каждый ответ снабжается коэффициентами уверенности:

– Cf1 – предназначен для идентификации ответа (верный ответ отмечается, неверный ответ отмечается 0);

– Cf2 – предназначен для идентификации сложности ответа (значение $Cf2 > 0$ означает увеличение сложности ответа при его истинности, а $Cf2 < 0$ означает уменьшение сложности ответа при его ложности). Коэффициенты влияют на оценки по тесту. Общий вид редактора БЗ представлен на рис. 5.

Каждый ответ можно снабдить изображением (формат bmp – редактируется и с помощью разноцветных дисков на рисунке обозначаются ответы).

Типы ответов на вопрос:

– 0 – с одним верным ответом;

– 1 – с множеством верных ответов;

– 2 – учитывается последовательность расположения ответов;

– 3 – свободно конструируемый ответ (пользователь вводит ответ с клавиатуры);

– 4 – выбор вариантов ответов (пользователю предлагается несколько столбцов ответов и он должен выбрать по одному верному ответу из каждого столбца). Оценка выставляется от выбранных правильных вариантов ответов;

– 5 – классифицировать ответы (пользователю предлагается ответы отнести к двум классам);

– 6 – список ответов (пользователю предлагается выбрать из выпадающего списка варианты ответов);

– 7 – ситуационная задача (при выборе правильного ответа задается уточняющий вопрос относительно выбранного ответа);

– 8 – несколько ответов из разных альтернатив (пользователю предлагается несколько столбцов ответов, и он может выбрать несколько верных ответов из каждого столбца). Оценка выставляется от выбранных правильных вариантов ответов.

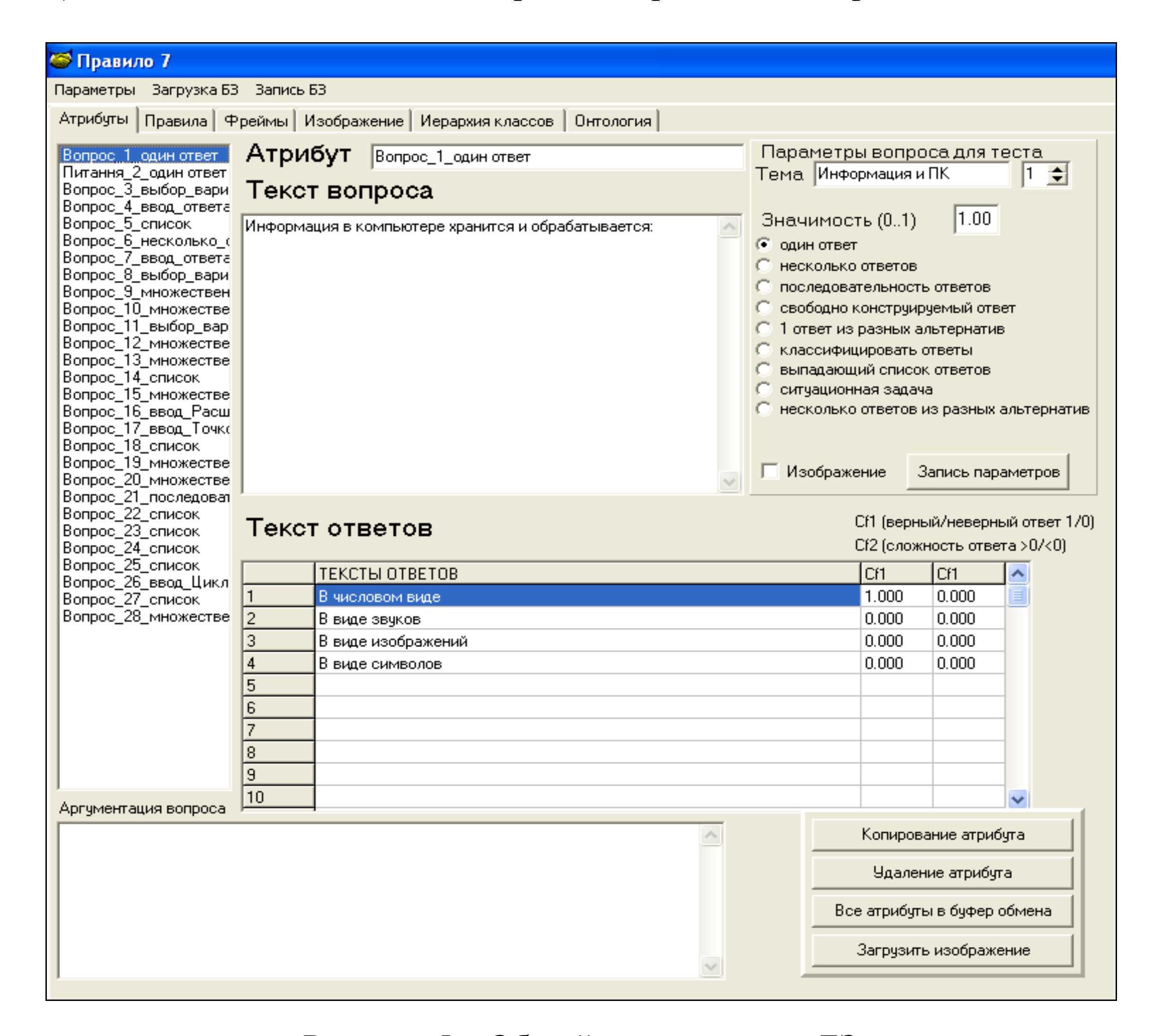


Рисунок 5 – Общий вид редактора БЗ

При выборе опции параметры – значимость вопроса и виды ответов в правом окне редактора появляется панель с параметрами вопроса.

Модель возмущения правил базы знаний

Пусть A = {a_1, a_2,..., a_n } – множество атрибутов предметной области, $D_{j\ t=0}$ – домен (множество возможных значений $a_j \in A$) в момент времени t = 0, $D_{j\ t=n}$ – домен (множество возможных значений $a_j \in A$) в момент времени t = n. Причем мощность домена может как увеличиваться, так и уменьшаться.

Атрибут, его значение и операция между ними определяют высказывание. Пусть Q = {q: A -> V} – множество высказываний.

На множестве Q определим две функции:

1. Функция $\mu : Q \to [0,\ldots,1]$, позволяет интерпретировать высказывание q с точки зрения его истинности.

2. Функция $\theta : Q \to [0,\ldots,1]$, позволяет интерпретировать высказывание q с точки зрения его значимости.

Тогда триплет f = (q, μ(q), θ(q)), где $q \in Q$, назовем фактом предметной области. По умолчанию можем считать, что факт $f_{t=0}$ образован в момент времени $t_0 = 0$. Если в момент времени $t = t_1$ выбирается другое значение из домена атрибута, обозначим f_{t1}, и во всех правилах БЗ происходит замена факта f_{t0} на f_{t1}, то такое действие над БЗ назовем возмущением ФС в момент времени $t = t_1$. Заметим, что возмущение узкое, поскольку при нем логическая структура ФС и последовательность расположения правил в БЗ не меняется.

Таким образом, получаем состояние ФС в момент времени $t = t_1$. Это состояние можно сохранить, записав его в отдельный файл, провести тестирование и проанализировать ход консультации на предмет содержательности БЗ в момент времени $t = t_1$. Если эксперт считает, что состояние БЗ в момент $t = t_1$ важно, то он может рекомендовать пользователю БЗ для эксплуатации. Если нет, то можно вызвать возмущение ФС в момент времени $t = t_2$ и повторить предыдущие шаги. Затем можно активизировать другой атрибут и проанализировать изменения БЗ. В результате таких специальных возмущений БЗ эволюционирует и эксперт может найти когнитивную ситуацию, которая не была им обнаружена при составлении БЗ. Такие возмущения БЗ используются на стадии ее отладки.

Определим ситуацию s_{t0} в момент времени $t = t_0$ как набор фактов, связанных между собой знаками конъюнкции, дизъюнкции или отрицанием. Например,

$$s_{t0} = f_{1\ t0}\ \&\ f_{2\ t0}\ U\ (\neg f_{3\ t0}).$$

Обозначим через S = {s} множество всех ситуаций предметной области и $p: S \to S$ некоторое преобразование множества S на себя. Рассмотрим итерации этого отображения, то есть результаты его многократного применения. Будем считать, что произвольная точка $s_0 \in S$ в момент времени $t = t_0$ за время $t_1 = 1$ переходит в точку $s_1 = p(s_0) \in S$. Тогда за время $t_2 = 2$ эта точка перейдет в точку $s_2 = p(s_1) = p(\ p(s_0))$ и т.д. В результате этих преобразований получаем траекторию эволюции состояния БЗ.

Пусть R = {r: S → S} – множество преобразований (правил) следующего вида:

r : ЕСЛИ <АНТЕЦЕДЕНТ> ТО <КОНСЕКВЕНТ>,

где <АНТЕЦЕДЕНТ> (условие – ситуация s_1), <КОНСЕКВЕНТ> (вывод, следствие – ситуация s_2). Если ситуация s_1 в правиле принимает значение истина, тогда ситуация s_2 получает значение истина и добавляется к S. Другими словами, если си-

туация $s_1 \in$ S в правиле принимает значение истина, тогда существует отображение q_j: A -> D, формирующее ситуацию s_2:

$$s_2 = q_1 * q_2 * q_n \text{ и * знак операций } \&, U, \neg.$$

Если ситуация s_1 в правиле принимает значение ложь, тогда ситуация s_2 не добавляется к S. Например, ЕСЛИ s ТО q_4. Здесь s – ситуация (3), q_4 – некоторое высказывание. Результатом выполнения правила будет создан новый факт в момент времени, например, t_3 = 3:

$f_{4\,t3} = (q_4, \mu(q_4), \theta(q_4))$,

$\mu(q_4) = k * \max [\min(\mu(q_1), \mu(q_2)), (1— \mu(q_3))]$.

$\theta(q_4) = k * \max [\min(\theta(q_1), \theta(q_2)), (1— \theta(q_3))]$.

Выводы

В данной работе рассмотрены алгоритмы построения различных моделей баз знаний для функциональной системы, реализованной в системе «КАРКАС»: модель базы знаний для определения риска возникновения ишемической болезни сердца, модель базы знаний для определения компетентности по информатике и модель возмущения правил базы знаний.

Установлено, что системы правил являются основным способом синтеза и представления множеств (планов) отношения на множестве объектов. Фактор, который упорядочивает объекты (частичный порядок), превращая ее в целенаправленную систему – это отображения r.

Классы составляют самоорганизующую совокупность для работы функциональной системы.

Частичный порядок объектов осуществляется в результате, когда консеквент одного правила, например p_1, содержится в антецеденте правила p_2, тогда объект 2 старше объекта 1.

Важная роль атрибутов заключается в том, что они задают отношение между объектами онтологии, например, значением атрибута является другой объект. Таким образом, порождается связь, вложенность объектов друг в друга.

Литература

1. Анохин П.К. Системные механизмы высшей нервной деятельности / Анохин П.К. – М. : Наука, 1979. – 453 с.
2. Бурдаєв В.П. Системи навчання з елементами штучного інтелекту / Бурдаєв В.П. – Харків : Вид. ХНЕУ, 2009. – 400 с.
3. Бурдаєв В.П. Моделі баз знань / Бурдаєв В.П. – Харків : Вид. ХНЕУ, 2010. – 300 с.
4. Бурдаев В.П. Модель функциональной системы динамической предметной области / В.П. Бурдаев // Искусственный интеллект. – 2011. – № 3. – С. 439-448.

Literatura

1. Anokhin P.K System mechanisms of the higher nervous activity. M: Science, 1979. - 453 s.
2. Burdayev V.P. Systems of learning with elements artificial intelligence Kharkiv: KNEU, 2009. – 400 s.
3. Burdayev V.P. Models of knowledge bases. Kharkiv: KNEU, 2010. – 300 s.
4. Burdayev V.P. Iskusstvennyj intellect. 2011. № 3. S. 439-448.

RESUME

V.P. Burdayev

Rules Formation of the Knowledge Base for Functional System

In the article, various models of knowledge bases for the functional system realized in system “KARKAS” are considered, i.e. model of the knowledge base for definition of risk of occurrence of an ischemic heart trouble, model of the knowledge base for competence definition on computer science and model of indignation of rules of the knowledge base.

The model ontology subject domain in system “KARKAS” consists of hierarchy of classes of subject domain, communications between them (conclusion rules) which operate within the limits of this model. In the system, the interpretation mechanism ontology in the conditions of dynamic change of its parameters (a base class, communications between classes and interactions of objects of classes) is offered. Technical realization for ontology is considered. The system is constructed by a modular principle and for this reason has possibility of connection of other additional modules. In architecture of system, it is possible to allocate the following basic modules: the loader; the module for working out knowledge base; the consultation module; the module cluster analysis of the data.

The subject domain model is considered as functional system, which the result makes organizing impact on all stages of formation ontology in. Classes and communications between it can be considered as a logic design of functional system.

In the system “KARKAS”, the functional system is the formalized reflexion of subject domain in the form of hierarchical structure of a set of managing directors of a component, which co-operate among themselves for overall objective achievement.

Статья поступила в редакцию 31.05.2012.